Foundations of Computing Projects 2026-27

Contents

Energy-Efficient Deep Learning with Sparse Subnetworks

Supervisor: Frederik Mallmann-Trenn

Areas: Artificial Intelligence, Machine learning / Deep learning, Foundations of computing (algorithms, computational complexity)

(Back to Scholarship Not Allocated)

Project Description

Supervisor: Dr Frederik Mallmann-Trenn
Host: Algorithm & Data Analysis (ADA) Group, King's College London
Theme: Energy-efficient machine learning, neural network sparsification, randomized processes
Keywords: supermasks / Edge-Popup, lottery-ticket hypothesis, sparsity, PyTorch/TensorFlow, energy measurement

Background:
Modern AI models are powerful—but they're also expensive in energy. That means higher cloud costs, bigger carbon footprints, and shorter battery life on phones and wearables. One promising route to cut energy is to run much smaller "sparse" networks that keep accuracy high while doing far less computation. This project studies a striking idea: you can often find a high-performing subnetwork inside a randomly initialised model by learning only a binary mask ("supermask")—no heavy weight training required. If we can understand and systematise this, we can make AI cheaper, greener, and more accessible.

Starting reference:
Ramanujan et al., What's Hidden in a Randomly Weighted Neural Network? (CVPR 2020)
https://openaccess.thecvf.com/content_CVPR_2020/html/Ramanujan_Whats_Hidden_in_a_Randomly_Weighted_Neural_Network_CVPR_2020_paper.html

Project overview

You will combine theory (probability and randomized processes) with hands-on engineering to reproduce, extend, and understand supermask methods (also called Edge-Popup).

Objectives:

1) Reproduce and clarify the baseline
2) Implement supermask/Edge-Popup for standard architectures (MLPs, CNNs).
3) Match key results from the paper; run careful ablations (how we parameterise masks, sparsity levels, initialisation, training schedules).
4) Make it practical and energy-aware
5) Explore structured sparsity (e.g., channel/block patterns) that hardware can exploit.
6) Explain when it works (theory)
7) Develop conditions for the existence of good sparse subnetworks under random initialisation (as functions of width, depth, and target sparsity k).
8) Study recoverability: when can simple procedures reliably find such subnetworks?
9) Map trade-offs (accuracy vs sparsity; compute vs expressivity), using tools from concentration of measure, random matrices/graphs, and randomized algorithms.

What you'll gain
1) Expertise in energy-efficient deep learning with reproducible engineering.
2) Deep understanding of randomized processes in modern ML.

Required background (must-have)

1) Mathematics: very solid probability/randomized processes (concentration inequalities, asymptotics), linear algebra, and optimisation.
2) Programming: strong PyTorch or TensorFlow skills (training loops, data pipelines, experiment tracking).
3) Ability to write clean, tested, reproducible research code (git; clear experiment logs).

References

Ramanujan et al., What's Hidden in a Randomly Weighted Neural Network? (CVPR 2020)
https://openaccess.thecvf.com/content_CVPR_2020/html/Ramanujan_Whats_Hidden_in_a_Randomly_Weighted_Neural_Network_CVPR_2020_paper.html

Implementing Differential Privacy in Neural Networks to Enhance Data Security and Anonymization

Supervisor: Frederik Mallmann-Trenn

Areas: Artificial Intelligence, Machine learning / Deep learning, Foundations of computing (algorithms, computational complexity)

(Back to Scholarship Not Allocated)

Project Description

Abstract: This PhD project aims to address the increasing need for robust privacy-preserving mechanisms in machine learning, particularly focusing on the application of differential privacy within neural networks. With the pervasive use of deep learning in processing sensitive information, there is a critical need to develop techniques that can protect individual data points from being reverse-engineered or identified. This research will explore innovative methods to integrate differential privacy into neural network architectures, ensuring the confidentiality of training datasets while maintaining the utility of the models.

Introduction: As neural networks become more ingrained in handling sensitive data, the potential for privacy breaches escalates. Differential privacy provides a framework to quantify and control the privacy loss incurred when releasing information about a dataset. This project will delve into the optimization of differential privacy in neural networks, balancing the trade-off between privacy protection and the predictive performance of the models.

Objectives: To conduct a comprehensive literature review on current approaches and challenges of applying differential privacy in neural networks.
To develop a theoretical framework for differential privacy that is specifically tailored to neural network applications.
To design, implement, and evaluate new algorithms that integrate differential privacy into neural network training processes without significantly degrading model accuracy.
To create a benchmark dataset and evaluation metrics for assessing the performance of privacy-preserving neural networks.
To investigate the impact of differential privacy on various neural network architectures and learning tasks, such as classification, regression, and generative models.
Methodology:
The project will utilize a combination of theoretical, experimental, and empirical methods. Initial efforts will focus on the theoretical underpinnings of differential privacy and its mathematical integration into neural network algorithms. Following this, experimental simulations using synthetic and real-world datasets will be conducted to assess the viability and performance of the proposed models. Empirical validation will be performed by comparing the new models with state-of-the-art privacy-preserving techniques.

It is absolutely necessary to have a strong math and stats background.

References

Dwork, C. and Roth, A., 2014. The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science, 9(3—4), pp.211-407.
https://www.cis.upenn.edu/~aaroth/Papers/privacybook.pdf
Abadi, M., Chu, A., Goodfellow, I., McMahan, H.B., Mironov, I., Talwar, K. and Zhang, L., 2016, October. Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC conference on computer and communications security (pp. 308-318).
https://arxiv.org/pdf/1607.00133.pdf

Software sustainability analysis and improvement

Supervisor: Kevin Lano

Areas: Machine learning / Deep learning, Systems (software engineering, programming), Foundations of computing (algorithms, computational complexity)

(Back to Scholarship Not Allocated)

Project Description

The project would consider techniques for analysing software sustainability (in the sense of energy use and energy efficiency) using either rule-based analysis and refactoring, or by the use of deep learning techniques such as LLMs to identify energy use flaws and potential refactorings.

Energy-efficiency improvement of machine learning systems is particularly important and could be the focus of the research. Equally, energy-efficiency improvement of mobile apps is another possible focus.

There is the potential for industrial collaboration in this area.

References

(Lano et al., 2024a) K. Lano et al., "Software modelling for sustainable software engineering", STAF 2024.
(Lano et al., 2024b) K. Lano et al., "Design Patterns for Software Sustainability", PLoP 2024
(Lano et al, 2025) K. Lano et al, "Sustainable Software Re-engineering", ECMFA 2025.

Mechanism Design for Robust AI Alignment

Supervisor: Carmine Ventre

Areas: Foundations of computing (algorithms, computational complexity), Artificial Intelligence

(Back to Scholarship Not Allocated)

Project Description

We propose to investigate a game-theoretic 'incentive engineering' solution to AI alignment using mechanism design. A dominant strategy incentive compatible (DSIC) mechanism would align AI systems to the designer's goals (captured by a so-called social choice function) since it would incentivise AI agents to act honestly, thereby implementing the function. This ideal solution relies on three problematic assumptions. It assumes AI agents are perfectly rational. It can only implement a limited class of functions due to known impossibility results.

The idea is to study novel solution concepts, inspired by the literature on behavioural economics. This will build on the landscape developed by the PI in his research on obviously dominant strategies. These strategies are played by agents with limited contingent reasoning skills or with access to data that is not granular enough to differentiate the payoff in each possible strategy profile. Concepts like OSP and SOSP restrict even further the class of social choice functions that are implementable vis-a-vis DSIC, by modelling agents who are only honest when it is obvious to them. NOM significantly relaxes DSIC, by assuming that agents will not manipulate the mechanism unless it is obvious to them. We will study mechanisms in the largely unexplored space between DSIC and NOM, thus guaranteeing alignment for AI agents that may find sophisticated manipulations non-obvious. We will test ideas developed at the interface of OSP and DSIC. To hope is to define a class of nested mechanisms that become more powerful as AI agents are less capable, making the results robust against advancing AI capabilities. This directly addresses the limitations above by guaranteeing alignment for agents that may refrain from some non-obvious manipulations.

Prospective applicants are encouraged to consult the publications of Prof Ventre
at https://kclpure.kcl.ac.uk/portal/en/persons/carmine.ventre/publications/.

Communication, Information, and Robustness in Trading Networks

Supervisor: Edwin Lock / Carmine Ventre

Areas: Machine learning / Deep learning, Foundations of computing (algorithms, computational complexity), Algorithmic Game Theory

(Back to Scholarship Not Allocated)

Project Description

Many markets function without a central auctioneer, instead relying on decentralised interactions between participants in environments characterised by bilateral negotiation and limited information exchange. This project will investigate the computational foundations of such decentralised trading. It will combine tools from algorithmic game theory, decentralised computing, and machine learning to explain how stable and equitable outcomes can arise, or fail to arise, when agents interact strategically in different informational environments.

This project focuses on how the amount and type of information exchanged between agents shapes outcomes in decentralised markets. Examples of such markets include financial markets and trading mechanisms implemented on the blockchain. In these settings, communication can be costly (e.g., blockchain transaction fees) or sensitive (e.g., a reluctance to reveal strategically valuable data). The project will investigate trade-offs between communication costs and the attainability of stable outcomes, as well as how incomplete information and strategic misrepresentation affect equilibrium behaviour. It will draw on tools from communication complexity and mechanism design, complemented by multi-agent reinforcement learning to simulate negotiation protocols under different informational constraints.

The student will gain expertise in algorithmic game theory, complexity analysis, market design, and machine learning. They will join a cross-disciplinary research environment at the intersection of computer science and economics, with potential opportunities for collaboration in both academia and industry (e.g. financial markets, commodity markets, and digital marketplaces). The project provides training in both rigorous mathematical analysis and computational experimentation, equipping the student with a broad skillset relevant to academia, industry, and policy.